A Portal for Rocket Scientists: Developing a New Internal Home Page for the Jet Propulsion Laboratory

Douglas Hughes[1], Jayne Dutra[2], and Jeanne Holm[3]

[1] Jet Propulsion Laboratory, Mail Stop 171-264, 4800 Oak Grove Drive
Pasadena, California 91109-8099 USA
Douglas.Hughes@jpl.nasa.gov
[2] Jet Propulsion Laboratory, Mail Stop 171-264, 4800 Oak Grove Drive
Pasadena, California 91109-8099 USA
Jayne.Dutra@jpl.nasa.gov
[3] Jet Propulsion Laboratory, Mail Stop 202-204, 4800 Oak Grove Drive
Pasadena, California 91109-8099 USA
Jeanne.Holm@jpl.nasa.gov

Abstract. In 1999, a small team at the Jet Propulsion Laboratory set out to create a prototype enterprise information portal for use throughout the National Aeronautics and Space Administration (NASA). May of 2001 saw the first instance of that portal in a limited evaluation release with very positive initial results. Increasing customer demand for new data channels and sub-portals bode well for the product, but ongoing funding and robust operational support remain unresolved issues. Follow JPL's Knowledge Management Navigation Team as they confront a bewildering market of rapidly evolving commercial software products, survive the dot.com body snatchers, gather and define sometimes conflicting customer requirements, blaze new trails through hostile IT jungles, create new paradigms in software development and deployment, survive reorganizations, and escape other perils of a large-scale, multi-vendor COTS integration.

1 Introduction

The Jet Propulsion Laboratory (JPL), managed by the California Institute of Technology (Caltech), is NASA's lead center for robotic exploration of the solar system. Within JPL, the Institutional Computing and Information Systems (ICIS) organization works to ensure that JPL has an information and computing environment that directly contributes to the success of JPL's missions. It is a tough job, considerably less "glamorous" than the Lab's core work of designing, building, and operating spacecraft. If your co-worker guides exotic robotic rovers on Mars for a living, it's really hard to get them excited about some killer COTS application…

However, the enterprise information portal may just do the trick, especially if it lives up to the promise of the technology. To gain a bit of background about JPL, we invite you to visit http://www.jpl.nasa.gov.

1.1 Background

In 1998, a small group at JPL was chartered to conduct a system engineering study of knowledge management and how it should be implemented. That team gathered requirements, spoke to customers and service providers, and created an architecture of services, processes, and system by which JPL could more effectively manage its intellectual capital and encourage knowledge sharing and reuse. Embodied in *A Knowledge Management Architecture for JPL,* was a specific charter to form a Knowledge Navigation team to

> *"Create an enterprise Web gateway to JPL's knowledge resources and easily customizable personal and group Web sites for easy access to institutional information and targeted delivery of information requested by individuals or workgroups."*

This involved both reengineering the existing Electronic Lab-wide Information Access Site (ELIAS) (JPL's internal home page) and creating a portal, complete with a robust search engine and institutional taxonomy covering all of the Lab's major repositories.

In early 1999, soon after JPL decided how to proceed in KM, IT engineers from NASA's Goddard Space Flight Center (GSFC), NASA's Langley Research Center (LaRC), and the Jet Propulsion Laboratory (JPL) answered a call for proposals from the NASA Chief Information Officer (CIO), Lee Holcomb. Together they developed a proposal for a *Federated Knowledge Management Architecture Prototype* for NASA. JPL had overall leadership of the knowledge management architecture, as well as for an element of the proposal entitled "Knowledge Portals".

Within months, the lead author for the portal proposal left for a dot.com startup, the co-author was promoted out of the loop, and the lead author on another proposal was transferred. Armed with courage, some funding, and little else, the KM Navigation team boldly entered the trackless COTS wilderness in search of a portal for rocket scientists.

2 Requirements Engineering

Perhaps the KM Navigation team's most valuable achievement was its fairly rigorous requirements engineering process that led to the publishing of four requirements documents: *JPL Internal Search Engine Requirements, JPL Portal System*

Requirements and JPL Portal Prototype System Requirements. The process followed was specifically tailored for COTS system development and delivery using the appropriate JPL internal, NASA, INCOSE, and IEEE documents as guidelines.

Just how did we get all those requirements? The KM Navigation team queried a wide range of the Lab's portal requirements sources, covering information consumers, repository owners and publishers, information service providers, and others. Led by the KM Navigation team's communications' experts, individuals, technical groups, line organizations, and projects were asked to contribute requirements and did so willingly. Interviews, focus groups, surveys, and outreach meetings were among the methods utilized. The final requirements impressed the team with their quality and quantity. The KM Navigation team also supplemented the user requirements with requirements that supported performance, hosting, interfaces, and security.[1]

The goals of the team were fairly lofty and aggressive, reflecting the desire to fully serve our customers and our naiveté about the lack of some essential Lab infrastructure that would prove disarming. In deference to the COTS portal and search engine market research we had conducted, and the relatively large size of the task, it was decided to completely "outsource" the portal implementation. This was January of 2000 (just before the stock market "crash" for e-business and Internet services providers) and we did not have a complete picture of the immaturity in the COTS, system integrator, and ASP markets.

3 Acquisition

In January 2000, the team issued a formal Request for Information (RFI) to 18 credible portal vendors and large-scale software integrators. The results received in the 90-day response period did not inspire our confidence, but did serve to refine our requirements and refresh our COTS market survey. On June 6, 2000, we entered the formal procurement process by issuing a Request for Proposal (RFP) to 38 portal vendors, large-scale integrators, and IT consulting groups. Once you waded past the procurement boilerplate and legal Teflon, the RFP was fairly simple. The small number of formal responses to the RFP received by discouraging in their implementation approach and very wide-ranging in cost. The process had failed to net a service provider capable of successfully implementing our requirements and whose costs remotely matched their offering.

---

[1] To the extent possible, we followed the definitions and advice about requirements contained in *Characteristics of Good Requirements*, by Pradip Kar and Michelle Bailey, presented at the 1996 INCOSE Symposium and prepared the Requirements Working Group of the International Council on Systems Engineering. This paper is publicly available at http://www.incose.org/rwg/goodreqs.html.

Given the RFP responses and our increased understanding of the COTS market for portal and search engine products, it was our conclusion that a combination of an iPlanet Portal Server (http://www.iplanet.com/products/iplanet_portal/home_portal.html), Autonomy search engine (http://www.autonomy.com/autonomy/dynamic/autopage10.shtml), and Compass spider (now a part of the iPlanet Portal server) would provide the basis for the portal. The relative maturity of the COTS software and the interest and support shown by the vendors were additional influences. These vendors were already on contract with NASA and we were able to procure initial software in August 2000. In the same time frame, an order for four supporting Sun servers and a Ciprico disk array was placed.

4 Architecture and Design

The collaboration of three software vendors is generally not a simple process and results may vary. Fortunately, iPlanet, Autonomy, and grapeVINE were self-motivated and friendly towards each other. Eventually, iPlanet absorbed grapeVINE. The Compass spider and search engine was integrated into the iPlanet Portal Server and the grapeVINE product itself became known as iPlanet Personalized Knowledge Services (PKS).

The combination of three major COTS products fused to make a portal is a bit unusual, but with the help of the vendors and their skilled professional services people, almost anything (you can afford) is possible. The initial portal programming was accomplished in a seven-week period utilizing an iPlanet Professional Services integrator. Autonomy consultants were utilized for installation, configuration, and programming.

The following figure is a high-level representation of the JPL portal architecture and system design.

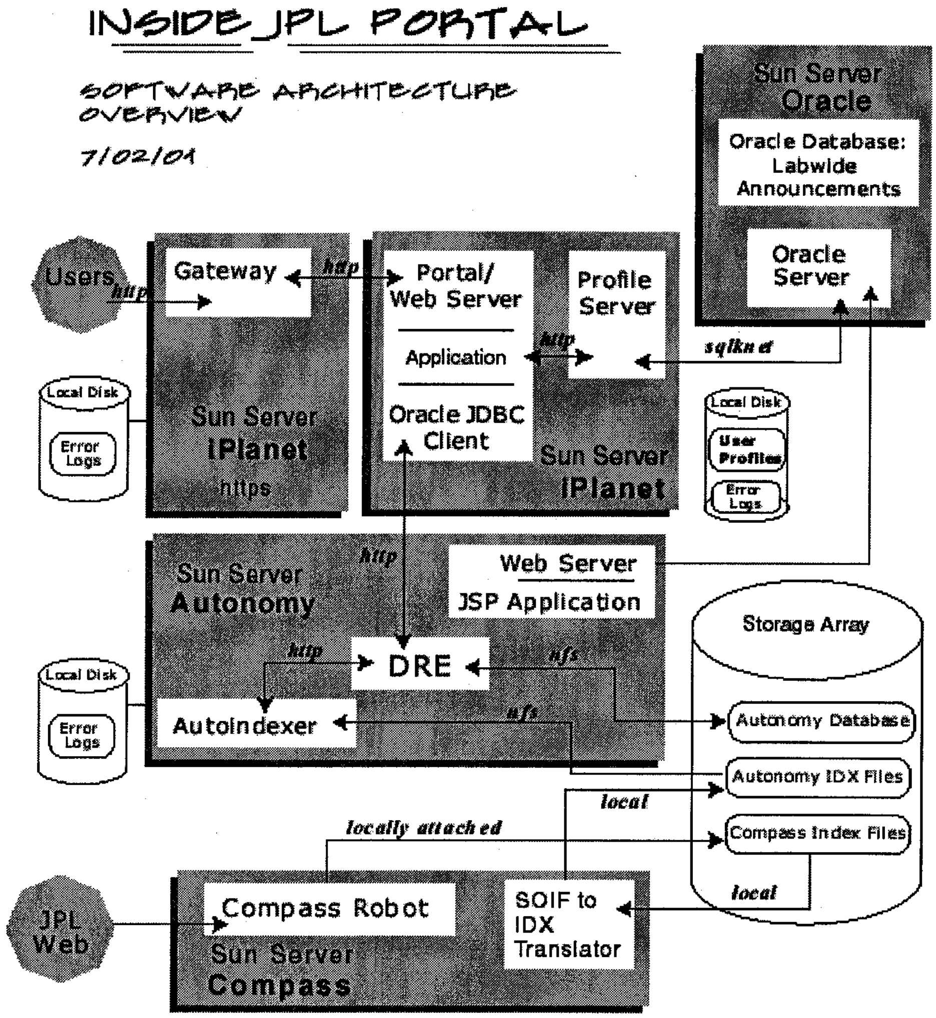

Fig. 1. High-Level Portal Architecture and System Design

The *JPL Portal System Requirements* included the user interface requirements. The implementation of the user interface in the iPlanet Portal Server framework manifested itself as 20 " data channels" displayed within a Web browser. The data channels can be grouped as follows:

Directory to JPL Web Space

- The centerpiece of the portal as it uses a carefully developed taxonomy to organize Web sites that are extensively used by JPL employees

Personal Information

- User Information - Allows users to update their name and greeting to personalize their content page.
- My Bookmarks – Allows the user to add their favorite internal/external bookmarks to their portal page.
- My Calendars – Contains links to one NASA and three JPL calendars.
- My Project Libraries – Contains links to desired internal JPL DocuShare (http://www.xerox.com/go/xrx/template/012.jsp?Xcntry=USA&Xlang=en_US&Xseg=corp&prodID=DocuShare) electronic libraries

Search

- Search JPL – Interface to the Autonomy search engine and the non-password protected information in JPL Web repositories and sites.
- Find a Person – Web interface into the JPL electronic phone book.
- Google Internet Search – Provided by Google (http://www.google.com) and directly accesses their Internet search engine and results.

General News

- Headline News: New York Times - Provided by iSyndicate (http://www.isyndicate.com), contains New York Times headline links and is updated daily.
- Headline News: Space.com – Provides information related to space and space-related content.
- Aviation Week – Contains *Aviation Week* (http://www.aviationnow.com) headline links and is updated daily.
- Weather and Traffic – Provides access to the services of weather.com (http://www.weather.com), the JPL weather station, and a graphical Los Angeles freeway traffic condition map from smartraveller.com (http://www.smartraveller.com).

Internal Content

- Lab-wide Announcements – Designed to replace Lab-wide e-mail and can be published to by any employee through a home-grown, single-purpose "content management" system.
- Institutional News – Contains links to the latest NASA and JPL press releases.
- Science Links – Contains internal and external links requested by the Lab's science community.
- Engineering Links - Contains internal and external links requested by the Lab's science community.
- Earth and Space Science Division (32) – Created to support this Division at JPL as a prototype for similar focused data channels.
- Business Links – Contains frequently used internal JPL business links
- Quick Links – Contains the eight most-used links.

- This Week – Contains the Lab's weekly bulletin, which is not to be confused with the Lab's biweekly printed newspaper, the *Universe,* or it's e-news site, the *Daily Planet.*

The following is a screen shot of the portal showing some of the data channels.

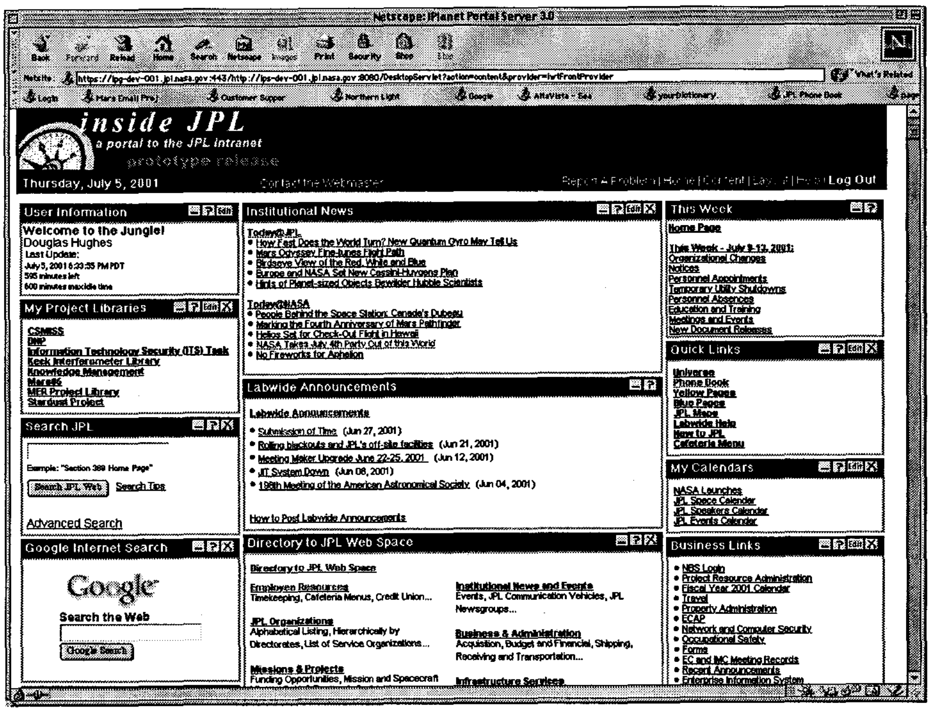

Fig. 2. Inside JPL – The Prototype Enterprise Intranet Portal for JPL

5 Software Development

Some of the more interesting and intricate software development in the system was the glueware created to interface the Compass spider (robot) and the Autonomy search engine. These two large COTS systems had never before been interfaced. Fortunately, the KM Navigation team had two sharp search experts and with some vendor support, the system works as expected. The following figure is a high-level representation of the Compass/Autonomy interface and data flow.

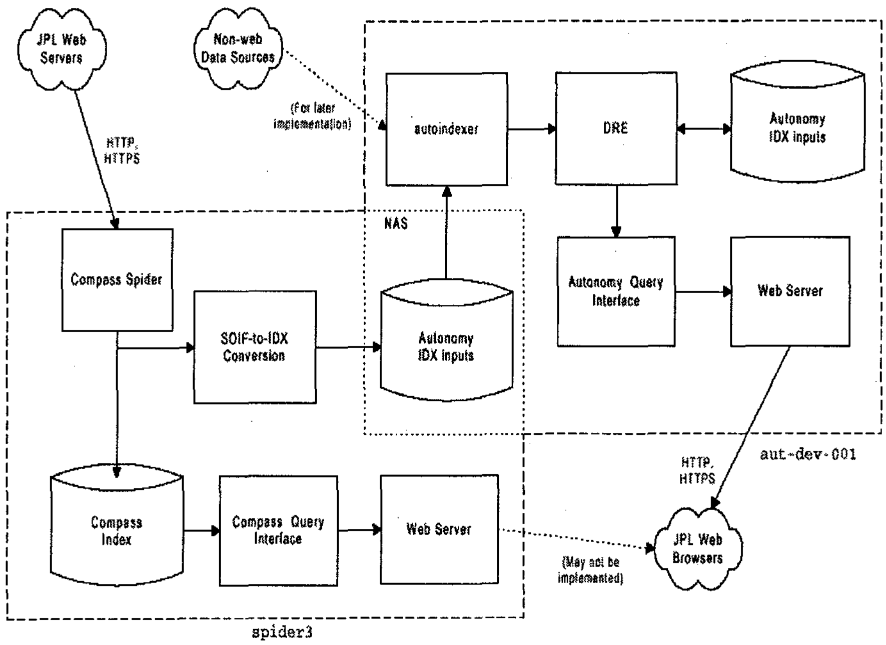

Fig. 3. Compass/Autonomy Interface

In order to shorten the delivery schedule and increase the quantity/quality of our initial portal offering, consultants from iPlanet and Autonomy were extensively utilized. This approach had the desired side effect of providing informal training to our staff without too much loss in consultant productivity. The burn rate for a major vendor's senior consultant is fairly significant, but the services are highly leveraged. An alternative was to send our internal staff to the vendor's product classes and contend with the schedule slip as the staff became more experienced.

Significant programming time was spent achieving a consistent " JPL look and feel." An online User Help Manual was created with Help screens made for each data channel. The search parameter input screen (both simple and advanced) and results pages took considerable programming. Because the search strategy incorporated a cross repository approach, the results pages included icons to identify the information source of each search hit. The search team also modified ranking to be especially meaningful to the JPL user.

The programming of the Compass spider and the Autonomy search engine interface have combined to give JPL portal users an order of magnitude increase in search index size and quality of search hits (increasing the search from 10,000 to 700,000 objects).
The full implementation of our portal has dependencies upon certain institutional infrastructure. Unfortunately, some of this supporting infrastructure is not yet

operational. JPL does not have a single source for authentication, which is not unusual, and IT system users need to remember a number of passwords. For the prototype period, the portal uses "self-registration and authentication." This will be abandoned once the institutional authentication mechanism becomes available. The Lab does not fully support a Web-based calendar system, so the portal was delivered with very limited calendar capability and access. The Lab does not support "instant messaging", so the portal is missing this important and requested feature. The Lab does support mail servers with IMAP protocol, but the portal team did not implement the NetMail feature of the iPlanet Portal Server due to resource and schedule constraints.

User-initiated publishing to the prototype portal is done through a home-grown interface, which only serves the Lab-wide Announcements data channel. It does contain some automated workflow and is supported by an Oracle database. This feature is augmented by a very helpful publishing guidelines page that reflects Lab policy and "readability" suggestions. There is a funded plan to investigate a COTS content management system in the next fiscal year.

6 (Some of) What We Learned

Lessons learned are emphasized at JPL, and the KM Navigation team has thoroughly documented and shared its lessons learned with people at JPL and at other NASA Centers. Here are some of our encapsulated/abbreviated lessons learned and general findings, in no particular order:

1. Investment in user engagement throughout the lifecycle pays large dividends, not only in product quality and acceptance, but also in the vital area of broad institutional support.
2. The availability of a portal "storefront" Web site helped welcome users and introduced the concept of a portal, its benefits, data channels, and FAQs. It also served as a link to user surveys.
3. The iPlanet Portal Server software is maturing with the help of its user base feedback and competitive market forces. Although a Portal Channel Wizard is provided, significant programming needed to be done utilizing both Java servlets and Perl scripts. APIs, completely documented, generally required the assistance of iPlanet technical support to be correctly implemented. The addition or modification of a data channel utilizing a Java servlet requires the re-setting of the server in order to make the changes available to the users. The speed of the portal's desktop presentation build depends greatly upon the immediate availability of the dynamic content in the data channels.
4. Have a simple interface control document with every possible content supplier to the data channels. Given the dynamic nature of most Intranets and certainly the Internet, broken links are to be expected. Use automated software to check links and report failures.

5. The iPlanet Professional Services consultant was exemplary during his seven-week engagement and has provided helpful periodic follow-up support via phone and e-mail. The iPlanet technical support has been thorough and accurate in response to queries and problem reports.
6. The Compass spider performed best using a depth-limited search across the JPL domain. The API for programming the Compass robot (spider) and its documentation are excellent. Their use of an open standard ASCII-based human readable format (Summary Object Interchange Format – SOIF) for their indexable content meant that we could use Open Source freeware Perl modules for reading it and converting to the Autonomy IDX format. Our initial experience with both the product and the company's support team has been excellent. Although we have run into some unexpected problems, the vendor's underlying programming team has been extremely responsive to our issues and suggestions.
7. Autonomy has been a bit of a mixed bag for the Knowledge Navigation Team. The Autonomy product was purchased with the belief that its full-text search functionality based on algorithmic patterns and word clustering would allow the user to take advantage of querying by concept as opposed to querying by keywords. This goal remains elusive at present due to some complicating factors, the most overriding of which was the time pressure of the delivery schedule. The installation procedures for each module are written primarily for the PC and not for the Unix environment. The installation of the Autonomy product was a long process with three visits from Autonomy personnel required in order to get the system working correctly. The underlying issue is that Autonomy was developed on the PC and the consultants sent out to JPL did not have very much experience with the product on a Unix server. Complete Unix documentation has not yet been provided. The parameters and values that should be passed to the Autonomy Server for getting the query results in the desired formats is not clearly explained within the Autonomy product documentation. The full-text nature of the query engine means that documents containing a great deal of text are ranked higher than web sites that typically contain a larger number of graphics and less verbal content. The service performed well under stress testing. From a user standpoint, searches entered by customers did not seem to take an inordinate amount of time to perform nor did the development team receive any complaints about the length of time a result set took to be formulated and displayed.
8. Automated test tools were a helpful component of the testing process for the portal. The tools used to test the portal were WinRunner, LoadRunner, and Topaz from Mercury Interactive (http://www.mercuryinteractive.com). Full integration, acceptance and load/performance test plans were developed and executed. Manual testing was also done with good results. Regression testing is currently done manually, but the tool scripts will be suitably augmented when funding and personnel become available.

7 Where We Plan to Go...

The initial 90-day evaluation period of the portal ended on June 15, 2001. Feedback from the Remedy problem reporting system and the on-line user survey is being analyzed at the time of this writing. The initial results are favorable. Approval and funding to proceed with the next phase in the development and deployment cycle has been secured for at least a minimal roll out of the portal. . Part of the challenge is that users were eager enough to use the portal, the Navigation team was asked to continue to allow access to the prototype portal to its existing users. This is not without risk and associated maintenance cost, as links break and external content may significantly change in format or vanish entirely. The real risk is that the portal could become " stale" because the team cannot add new data channels, procure additional syndicated content, or create requested sub-portals in its current operating mode.

The real work is just beginning on the NASA-wide portal. It is expected that the developers will exploit the knowledge and experienced gained in implementing the JPL prototype portal in developing the NASA portal. For a job of this size, complexity, and risk, it is reasonable to both re-visit the portal market and to examine the use of a major systems integrator as a possible partner.

References

1. Doane, J., Hess, S., Cooper, L., Holm, J., Fuhrman, D., U'Ren, J.: A Knowledge Management Architecture for JPL. Jet Propulsion Laboratory (1999)
2. Sorensen, S., Renfrow, T., NASA Chief Information Office Information Technology Proposal, Federated Knowledge Management Architecture Prototype, Knowledge Navigation as a Component of a Knowledge Management Architecture. Jet Propulsion Laboratory (1999)

The research described in this paper was conducted at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.